

Regulating Autonomous Weapons: Recent Diplomatic & Political Developments

Peter M. Asaro, PhD

asarop@newschool.edu

Associate Professor

[School of Media Studies, The New School](#)

Affiliate Scholar

[Center for Internet and Society, Stanford Law School](#)

Co-Founder & Vice-Chair

[International Committee for Robot Arms Control](#)

Autonomous Weapons

- Brief History of the Campaign
- History of CCW Process
 - Defining Autonomous Weapons Systems (AWS)
 - Meaningful Human Control
- Pivot to UNGA Discussions
 - Parallels to TPNW
 - Relation to REAIM & Political Declaration
- Concepts: ICRC Recommendations (2-Tiered Approach)
 - Human Dignity is the Moral Foundation for Both
 - Challenges for Anti-Personnel & Non-IHL Uses

- Sept. 2009: ICRAC Founded
- April 2013: Stop Killer Robots Campaign Launched (UN HRC)
- Spring 2014-2016: UN CCW Experts Meeting

- 2017-2026: CCW Group of Governmental Experts Meetings
- 2023-2024: Regional & Global Meetings (Costa Rica, Luxembourg, Philippines, Sierra Leone, Austria)
- 2024-2025: UNGA 1st Comm (UNSG Report, UNGA Informals)



Coalition of NGOs

- 9 NGOs on Steering Committee
- 280+ NGO members
- 70+ Countries
- Putting pressure on their Governments to take action at the UN
- Raising Public Awareness of the Issue
- Informing Media

Different than Other Campaigns

- Not Specific Weapon, Control of All Weapons
- Pre-Emptive Arms Control
 - Getting Ahead of Technology
 - “Killer Robots”
- No Victims
- Shadows of Drones, Cyber, TPNW

Support from States

- 120+ Countries support Regulation
- 70+ Countries Support Negotiating a Legally Binding Treaty
- Positions
 - US, UK, Russia, Israel [Do Nothing/Non-Binding]
 - China?
 - Japan, South Korea, NATO countries
 - 33 Latin American +3 EU Countries sign Belen Communique
 - ECOWAS & Freetown Communique
 - 164 & 161 States Support UNGA Resolutions

Current CCW Situation

- GGE Mandate through 2026
 - Meeting of High Contracting Parties
- Elements of Instrument

“For the purposes of the work of the GGE and with the understanding that nothing has reached consensus until everything has reached consensus, the Group provisionally found consensus on the following formulations for the purpose of advancing its work on a set of elements of an instrument, without prejudging its nature, and other possible measures, in accordance with the Group’s mandate.”
- Chair’s “Rolling Text”
- Avoid Parallel, Distinct from AI in Military

Campaign Goals

- Legally Binding Treaty
 - Establish Clear Norms
- Move Discussion Beyond CCW
- Build Consensus Among States
 - Representing Civil Society
- Preserve Key Elements
 - Protection of Civilians
 - Human Rights & Human Dignity

Defining Autonomous Weapon Systems

- Lethal Autonomous Weapons Systems (LAWS)
- Fully Autonomous Weapons
- Killer Robots
- Humans-in-the-loop
- Decision Cycle/Kill Chain (USAF)
 - Find
 - Fix
 - Track
 - Target
 - Engage
 - Assess



Chair's Rolling Text

1. Within the scope of the application of the CCW, a lethal autonomous weapon system can be characterized as an integrated combination of one or more weapons and technological components, that can select and engage a target, without intervention by a human user in the execution of these tasks.
2. For the purposes of these elements of an instrument, lethal refers to the capacity to cause death to a person. The fact that a LAWS can be used in circumstances that do not result in death, such as to damage or destroy objects or to cause injury, does not exclude it from this characterization.
3. The above description is without prejudice to any future understanding and the potential modification of this characterization, as well as the possible exclusion of certain types of systems.

Reasons to Regulate

- Lowering Thresholds of Conflict (Cheaper, Less Risk)
- Arms Races & Regional/Global Instability
- Unintended Conflicts & Escalation of Conflict
- Vulnerable to Hacking & Cyberattack
 - AI Unreliability
- Rapid Proliferation (Including Non-State Actors)
- Accountability Gap (Undermining Law)
- **Dangers to Civilians & Threats to Human Rights (Legal & Ethical)**

Technology Drivers for Autonomy

- Heavy Investments in R&D
- Lower Risks to Own Soldiers
- Lower Personnel Requirements
- Ever Increasing Pace of Battle/Decision Cycle
- Ability to Operate Without Comm Link
- AI Hype (Super-Human promise)

Meaningful Human Control

- Origins & Emergent Norms
- “Requirements for Human-Machine Interaction”
- Appropriate Levels of Human Judgement (US DoD 3000.09)
- Chair’s Rolling Text
 - “Context-Appropriate Human Judgement and Control”
- Empty Vessel vs Negotiated Definition
 - St. Petersburg Declaration “Unnecessary Suffering & Superfluous Injury”

Recent Developments

- 2020: COVID
- 2021: ICRC Position Paper (Two-Tier Approach)
- 2022: Conflict in Ukraine, ChatGPT Launched
- 2023: Conflict in Gaza
- 2023: UNSG & ICRC President Joint Statement
 - Call for Treaty by 2026
- 2024: Vienna Meeting & UNSG Report
- 2025: UNGA Informal Meeting

Current Conflicts

- Ukraine
 - Drones & Autonomy
- Gaza
 - Reports of Autonomous Drones
 - AI Targeting: Decision Support Systems

Prohibited Systems (Tier 1) 2021

With a view to supporting current efforts to establish international limits on autonomous weapon systems that address the risks they raise, **the ICRC recommends that States adopt new legally binding rules.** In particular:

1. **Unpredictable autonomous weapon systems should be expressly ruled out**, notably because of their indiscriminate effects. This would best be achieved with a prohibition on autonomous weapon systems that are designed or used in a manner such that their effects cannot be sufficiently understood, predicted and explained.
2. In light of ethical considerations to safeguard humanity, and to uphold international humanitarian law rules for the protection of civilians and combatants hors de combat, **use of autonomous weapon systems to target human beings should be ruled out.** This would best be achieved through a prohibition on autonomous weapon systems that are designed or used to apply force against persons.

Regulated Systems (Tier 2)

1. In order to protect civilians and civilian objects, uphold the rules of international humanitarian law and safeguard humanity, **the design and use of autonomous weapon systems that would not be prohibited should be regulated**, including through a combination of:
 - **limits on the types of target**, such as constraining them to objects that are **military objectives by nature**
 - **limits on the duration, geographical scope and scale of use**, including to enable human judgement and control in relation to a specific attack
 - **limits on situations of use**, such as constraining them to situations where civilians or civilian objects are not present
 - **requirements for human-machine interaction**, notably to ensure effective human supervision, and timely intervention and deactivation.

Targeting Military Objectives

- Targeting Civilian Objects Should be Prohibited
 - Even if being used for Military Purpose
 - Requires Proportionality
- Occupied Military Vehicles & Structures?
- Are Uniformed Armed Combatants Military Objectives?

Jus in bello & Geneva Conventions

- Discrimination
 - Principle of Distinction
 - Combatants and Non-combatants
 - Hors de Combat (Wounded, Surrender, Shipwrecked)
 - Partisans (ICRC guidelines Direct Participation in Conflict)
 - Doctrine of Double-Effect
 - Intentions
 - Civilian Casualties & Infrastructure
- Proportionality
 - Response (Escalation, Diplomacy)
 - Risk of Civilian Harms
 - Principle of Least Harm (Unnecessary/Superfluous)
- Military Necessity / Principle of Precaution

Justified Killing in War

- It Matters *How* One is Killed
 - Justification (Self-Defense)
 - Not Arbitrary, Summary, Extra-Judicial
- Duty of Commanders
 - Take all Reasonable Precautions to Prevent Harms to Civilians & Civilian Infrastructure
 - Assess Military Necessity
 - For Each Attack

Improper Delegation

- AWS are Not Legal Agents
- AWS are Not Moral Agents
 - It is immoral (abrogation) to delegate to machines/algorithms
- Undermines Responsibility
 - Psychology of Human Decisions
 - Legal Accountability

Human Rights & Dignity

- Right To Life
- Human Right to Dignity
 - Prior to Right to Life
 - Beyond the Right to Life
 - Desecration
 - Dehumanization is Collective
 - Slavery & Torture
 - **Right to Have Rights (Arendt)**
 - **Right to Be Recognized/Respected as Human**

Proportionality & Dignity

- Cannot Pre-Judge/Pre-Authorize
 - Context matters, Timing Matters
 - German Constitutional Law Decision
 - No set numbers for acceptable harms
- Equality under Law/Moral Consideration
 - No Civilian is More Equal (morally)
 - Trolley Problems in Self-Driving Cars

Digital Dehumanization

- Permitting Automated Lethal Decision
 - General/Collective Dehumanization
- Reducing Humans to Digital Representations (Target Profiles)
 - For Fundamental Rights Claims (Life/Dignity)
 - Rights of Appeal/Explanation
- Other Claims Against Algorithmic Decisions?

Humans vs Military Objects

- **limits on the types of target**, such as constraining them to objects that are military objectives by nature
- **limits on situations of use**, such as constraining them to situations where civilians or civilian objects are not present
- Should not create Digital Representations of Humans for purpose of killing
 - OK to avoid causing collateral harms?
- Algorithms should not be designed to make Proportionality Judgements
 - Requires establishing/learning quantitative or comparative values of human lives, thus fundamentally dehumanizing
- AWS Permissible but Regulated when not targeting people, or only creating targeting profiles for “military objectives by nature”

CCW Outcomes & Next Steps

- Chair's Rolling Text
 - Consensus Document
- Consensus Rule
- Not Likely to be Legally Binding
- Inherent CCW Limitations
 - IHL not Human Rights
 - Policing & Border Control
 - Anti-Personnel Prohibitions
- Treaty Process: UNGA?

Thank You!

- Email: asarop@newschool.edu
- WWW: www.peterasaro.org